Org Divers Evol (2013) 13:77–85
DOI 10.1007/s13127-012-0102-6

ORIGINAL ARTICLE

ORGANISMS DIVERSITY & EVOLUTION

# Conservation genetics and phylogeography of the poorly known Middle Eastern terrapin *Mauremys caspica* (Testudines: Geoemydidae)

**Melita Vamberger · Heiko Stuckas · Dinçer Ayaz · Eva Graciá · Abdulhadi A. Aloufi · Johannes Els · Lyudmila F. Mazanaeva · Haji Gholi Kami · Uwe Fritz**

Received: 16 February 2012 / Accepted: 14 June 2012 / Published online: 14 July 2012


**Abstract** The West Asian stripe-necked terrapin *Mauremys caspica* is widespread throughout the Middle East—a region for which only few phylogeographic studies are available. Due to landscape alteration, pollution and intensification of water management, *M. caspica* is increasingly threatened. However, genetic diversity among and within populations is poorly known, impeding the identification of management units. Using a nearly rangewide sampling, we analyzed 14 microsatellite loci and mtDNA sequences in order to gain insight into the population structure and history of *M. caspica*. In agreement with a previous study, we found two clusters of mitochondrial haplotypes, with one cluster distributed in the east and the other in the west of the range. However, our microsatellite data suggested a more pronounced geographical structuring. When null alleles were coded as recessive with STRUCTURE 2.3.2, three clusters were revealed, with one cluster matching roughly the range of the western mitochondrial cluster, and the composite ranges of the two other microsatellite clusters correspond to the distribution of the eastern mitochondrial cluster. Naïve STRUCTURE analyses without correction for null alleles were congruent with respect to the two eastern microsatellite clusters, but subdivided the western cluster into two units, with an additional geographical divide corresponding to the 'Anatolian diagonal'—a well-known high mountain barrier impeding exchange between western and eastern taxa. In naïve analyses, the westernmost microsatellite cluster (from Central Anatolia) is quite isolated from the others, and its distinctness is also supported by fixation indices resembling the values among the other three clusters. One of the two eastern clusters is distributed in the Caucasus region plus Iran, and terrapins from Saudi Arabia and Bahrain constitute the second eastern cluster, supporting

**Electronic supplementary material** The online version of this article (doi:10.1007/s13127-012-0102-6) contains supplementary material, which is available to authorized users.

M. Vamberger · H. Stuckas · U. Fritz (✉)
Museum of Zoology (Museum für Tierkunde),
Senckenberg Dresden, A.B. Meyer Building,
01109 Dresden, Germany
e-mail: uwe.fritz@senckenberg.de

D. Ayaz
Department of Biology, Faculty of Science, Ege University,
Bornova-Izmir, Turkey

E. Graciá
Ecology Area, Department of Applied Biology,
Miguel Hernández University,
Av. de la Universidad, Torreblanca,
03202 Elche, Spain

A. A. Aloufi
Tabuk University,
P.O. Box 741, Tabuk 71491, Kingdom of Saudi Arabia

J. Els
Environment and Protected Areas Authority,
Breeding Centre for Endangered Arabian Wildlife,
P.O. Box 29922, Sharjah, United Arab Emirates

L. F. Mazanaeva
Department of Zoology, Dagestan State University,
Makhachkala,
Dagestan, Russia

H. G. Kami
Department of Biology, Faculty of Sciences, Golestan University,
Gorgan, Golestan Province, Iran

the view that these endangered populations are native. Coalescent-based analyses of our microsatellite data reveal for all four clusters bottlenecks 4,000–20,000 years ago, suggesting that climatic fluctuations of the Late Pleistocene and Holocene played an important role in shaping current genetic diversity. We propose that each of the four identified clusters, including the Central Anatolian one, should be treated as a distinct management unit. The presence of non-native terrapins in the animal trade of Bahrain highlights the danger of genetic pollution of the endangered Arabian populations. Further sampling is needed to elucidate the situation in southern and central Iran and Iraq. Our results confirm that genetic data do not support the validity of any of the three morphologically defined subspecies of *M. caspica*, and we propose that their usage be abandoned.

**Keywords** Reptilia · Divergence dating · Endangered species · Population structure · Subspecies

## Introduction

The stripe-necked terrapin *Mauremys caspica* (Gmelin 1774) is a medium-sized freshwater turtle that is widespread throughout the Middle East. It reaches a maximum shell length of approximately 25 cm and is distributed from Central Anatolia east- and southeastwards across Syria and the Caucasus Region to Iraq and Iran; isolated relict populations are known from Bahrain and adjacent Saudi Arabia (Anderson 1979; Fritz and Havaš 2007; Fritz and Wischuf 1997; Gasperetti et al. 1993). While the species is still common in many parts of its range, landscape alteration, pollution and intensification of water management in Turkey (Bayazit and Avci 1997; Harmancioğlu et al. 2001), Syria, Iraq (Beaumont 1998; Garstecki and Amr 2011) and Iran (Hashemi et al. 2012) are increasingly threatening the survival of many populations (personal observation; cf. also Biricik and Turğa 2011; Kinzelbach 1986). Of special conservational concern are the Arabian populations (Anon. 2003) that are confined to Bahrain, and the oasis complexes of Al Qatif, Al Hufhuf and Al Uqayr, Saudi Arabia (Gasperetti et al. 1993). There, the species faces the additional threat of the rapid urbanization of its habitat and unsustainable use of freshwater. However, Gasperetti et al. (1993) raised the question of whether these isolated Arabian populations might result from ancient introductions.

Until now, population structuring of *M. caspica* is poorly known, impeding the identification of management units (i.e., demographically largely independent populations with low exchange rates, Palsbøll et al. 2011). A first phylogeographic investigation using the mitochondrial cytochrome *b* (cyt *b*) and the nuclear C-mos genes found only weak differentiation, with two mitochondrial haplotype clusters, one occurring in the east and the other in the west of the range (Fritz et al. 2008). This study did not shed light on the possible non-native status of the Arabian populations, even though a high diversity of mitochondrial haplotypes was recorded for Bahrain.

In the present paper, we re-examine the phylogeography of *M. caspica* using a nearly rangewide sampling. Besides the mitochondrial cyt *b* gene, we use 14 rapidly evolving microsatellite loci (Vamberger et al. 2011) to achieve a better understanding of the phylogeography and population structuring of *M. caspica*. In particular, we re-assess the status of the critically endangered Arabian populations. Our results are intended to serve as a means for future conservation projects and to contribute to the scarce knowledge of the phylogeography of Middle Eastern biota.

## Materials and methods

### Laboratory procedures

Saliva, tissue and blood samples of 124 stripe-necked terrapins from 37 localities were studied, covering most of the species' range except Iraq (Table S1). Samples were preserved in ethanol or EDTA buffer (Arctander 1988) and stored at –80 °C in the tissue sample collection of the Museum of Zoology, Senckenberg Dresden, until processing.

DNA was isolated using the InnuPREP DNA Mini Kit (Analytik Jena, Jena, Germany) or the NucleoSpin Tissue Kit (Macherey-Nagel, Düren, Germany). For each sample, 14 unlinked microsatellite loci as characterized by Vamberger et al. (2011) were studied. Four or five loci each were combined in multiplex PCRs. The final volume of each multiplex PCR was 10 µl containing 0.5 U *Taq* polymerase (Bioron, Ludwigshafen, Germany) with the buffer recommended by the supplier and a final concentration of 1.5 mM $MgCl_2$ (Bioron), 0.2 mM of each dNTP (Fermentas, St. Leon-Rot, Germany), 0.25 µg bovine serum albumin (Fermentas), approximately 10–20 ng of total DNA, and different concentrations of primers, ranging from 0.05 µM to 0.7 µM. The forward primers were fluorescent-labelled. The PCR cycling conditions were as follows: 35 cycles with denaturation at 95 °C for 30 s but for 15 min for the first cycle, annealing at 56 °C for 30 s, and extension at 72 °C for 30 s but 30 min for the final cycle. PCR products were diluted with water in a ratio of 1:100. Fragment lengths were determined using an ABI 3130xl Genetic Analyser, the GeneScan-600 LIZ Size Standard and the software GENEMAPPER (Applied Biosystems, Foster City, CA).

In addition, the mitochondrial cyt *b* gene was sequenced in cases where the data for the same samples were not available from a previous study (Fritz et al. 2008). The respective mtDNA fragment was amplified for 43 samples

(Table S1) using the primer pair CytbG (Spinks et al. 2004) and mt-f-na (Fritz et al. 2006). PCR was performed in a final volume of 20 μl using 1 U *Taq* polymerase (Bioron) with the buffer recommended by the supplier and a final concentration of 0.2 mM of each dNTP (Fermentas), 0.5 μM of each primer and ca. 10–40 ng total DNA. The PCR cycling conditions were as follows: 35 cycles with denaturation at 95 °C for 45 s but 5 min for the first cycle, annealing at 58 °C for 45 s, and extension at 72 °C for 1 min but 10 min for the final cycle. PCR products were purified with the ExoSAP-IT enzymatic cleanup (USB Europe, Staufen, Germany) and sequenced using the forward primer mt-c-For2 (Fritz et al. 2006) and the newly designed reverse primer RI-neu (5'-GTG AAG TTG TCT GGG TCT CCT AG-3') on an ABI 3130xl using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). The resulting mtDNA fragments had a length of approximately 1,040 bp. Individual sequences were aligned in BIOEDIT 7.0.5.3 (Hall 1999), collapsed manually into haplotypes and merged with previously published data (see Fritz et al. 2008 for GenBank accession numbers).

For one previously published cyt *b* sequence corresponding to a distinct haplotype (Fritz et al. 2008), no microsatellite data could be produced because the sample was used up. However, its cyt *b* sequence was included in the respective analyses of mtDNA (Table S1).

## Data analyses

To ass'ess population structure, our microsatellite data were subjected to cluster analysis using STRUCTURE 2.3.2 (Falush et al. 2003, 2007; Hubisz et al. 2009; Pritchard et al. 2000), without a priori information about sampling sites. The main criterion for cluster delineation is the search for groups in Hardy-Weinberg equilibrium and linkage equilibrium. The search is conducted for each locus separately, which allows admixtures to be detected. The optimal number of clusters was revealed using posterior probabilities [highest lnP(D)] and the Δ*K* method of Evanno et al. (2005). Considering the number of distinct collection sites (37), the population structure was modelled using an upper bound of 40 and an admixture scenario with allele frequencies correlated, allowing the individuals to have mixed ancestries. The burn-in was set to $25\times10^4$ and the number of further MCMC runs to $75\times10^4$. Calculations were repeated ten times for each *K* ranging from 1 to 40; convergence of likelihood values was reached after the burn-in. Clustering results and individual admixture were visualized using bar plots. Individuals below a threshold of 80 % for cluster membership (Randi 2008) were treated as having mixed ancestries. Each STRUCTURE cluster and populations with sufficient individual numbers were tested for null alleles using MICRO-CHECKER 2.2.3 (van Oosterhout et al. 2004). Since null alleles were detected for some loci, the data set was re-analysed in STRUCTURE by coding null alleles as recessive (Falush et al. 2007) and otherwise using the same settings as above.

The demographic history of *M. caspica* was examined using MSVAR 1.3 (Beaumont 1999; Storz and Beaumont 2002). This software is based on a Bayesian coalescent method and allows historical changes in effective population size of the studied population to be inferred using observed allele distributions and frequencies under the assumption that a stable population with a determinate size ($N_1$) started to decrease or increase linearly or exponentially some time ago ($T$) towards the current population size ($N_0$). Based on this premise and using the Markov chain Monte Carlo approach, MSVAR estimates $N_0$, $N_1$ and $T$. Mutation rates ($\theta=2N_0\mu$) are assumed under the stepwise mutation model and, as the other parameters, are inferred separately across loci using priors with lognormal distributions. For inferring founder and bottleneck processes, the exponential growth model was applied. Based on the estimates of Bannikov (1951), generation time was set to 10 and 15 years. To examine the stability of results, the program was run four times with $2\times10^4$ iterations using the two generation times and different ranges of uninformative priors and hyperpriors (Table S2). The first 10 % of all iterations were discarded to avoid bias in parameter estimation due to starting conditions. The distribution of the remaining data was plotted against prior distributions to evaluate the consistency among different runs and to obtain the lower (10 %), the median (50 %) and upper quantile (90 %) of the posterior distributions. Given that genetic structure can generate false bottleneck signals (Chikhi et al. 2010), the data set was divided and analyses run independently for each of the four STRUCTURE clusters.

Mutational relationships of mitochondrial sequences were examined by parsimony network analysis as implemented in TCS 1.21 (Clement et al. 2000).

Based on the STRUCTURE clusters, diversity and divergence parameters were estimated for microsatellite and mitochondrial haplotype data. To compare the number and size of microsatellite alleles, a frequency table was produced using CONVERT 1.31 (Glaubitz 2004). Locus-specific observed ($H_O$) and expected heterozygosities ($H_E$) were estimated in ARLEQUIN 3.11 (Excoffier et al. 2005) and the same software was used to perform a locus-by-locus analysis of molecular variance (AMOVA; 10,000 permutations). Deviations from Hardy-Weinberg equilibrium were assessed using GENEPOP 4.0 (Rousset 2008), and the same software was used for determining locus-specific excess or deficiency of heterozygotes as expressed by the inbreeding coefficient $F_{IS}$ (Weir and Cockerham 1984). Statistical significance of $F_{IS}$ values was tested using FSTAT 2.9.3.2 (Goudet 1995). Exact *P* values for these inbreeding coefficients were estimated by the Markov Chain method (1,000 dememorization steps followed by

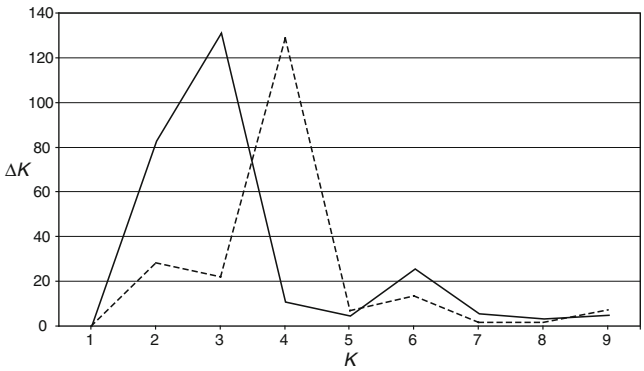


**Fig. 1** $\Delta K$ values for $K$ ranging from 1 to 9, showing the modal values for the runs corrected for null alleles (*solid line*) and the naïve runs (*dashed line*) of STRUCTURE

100,000 iterations) and resulting values were again Bonferroni-corrected. Finally, values for locus-specific allelic richness were calculated with the software FSTAT. Frequencies of mitochondrial haplotypes were obtained with ARLEQUIN and then used for an AMOVA (10,000 permutations). Mitochondrial nucleotide and haplotype diversities within STRUCTURE clusters were determined in DNASP 5.0 (Librado and Rozas 2009).

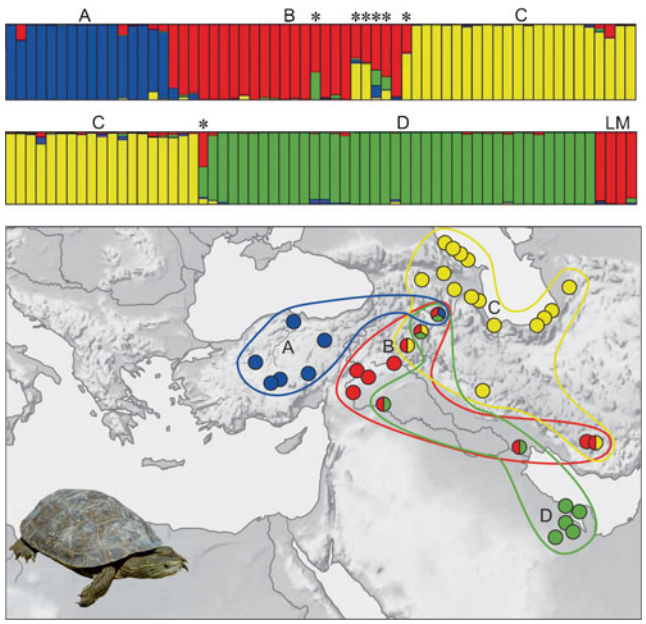


**Fig. 2** Population structuring in *Mauremys caspica* for $K$=4 (*top*) from the STRUCTURE run with the highest probability value. Within each cluster, an individual terrapin corresponds to one column that reflects its inferred ancestry. Individuals with mixed ancestries (below a threshold of 80 % for cluster membership) are indicated by *asterisks*. *Colours* of sampling sites in the map (*bottom*) correspond to STRUCTURE clusters; *slices* indicate presence of individuals with mixed ancestries of respective clusters. Mixed ancestries are also indicated by *shared subsets* of lines encircling symbols. *Upper-case letters* refer to clusters; *LM* to terrapins from the local market of Bahrain. When null alleles are coded as recessive, clusters *A* and *B* are lumped together and two individuals with mixed ancestry are referred to this more inclusive cluster (cf. Table S1)

## Results

### Population structuring

#### *Microsatellites*

Length polymorphisms at 14 microsatellite loci of 124 terrapins from 37 sites were subjected to Bayesian cluster analyses in STRUCTURE 2.3.2 using two different approaches. In a first naïve approach, the raw data were not corrected for null alleles and, in the second approach, null alleles were coded as recessive as suggested by Falush et al. (2007). The results of both methods were largely concordant and suggested a pronounced population structuring.

For the dataset without correction for null alleles, the highest mean posterior probability was obtained for $K$=9, and for the dataset with null alleles coded as recessive, $K$=6. However, the $\Delta K$ method of Evanno et al. (2005) revealed clearly lower values. Then, for the uncorrected data set the optimal number of clusters was four, whereas three was suggested as optimal for the data set corrected for null alleles (Fig. 1). In the following, we present and use the results of the $\Delta K$ method, because it is known that the

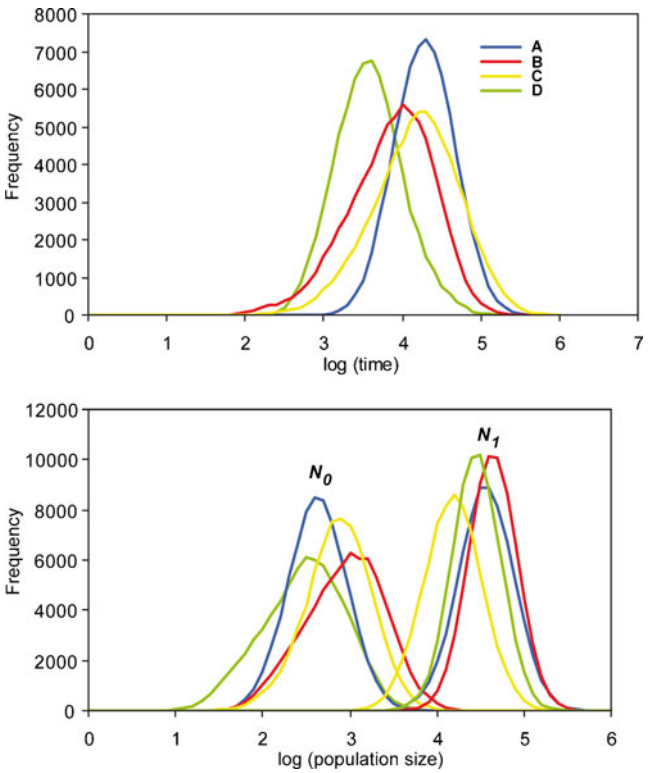


**Fig. 3** Posterior distributions of population demographic parameters for *Mauremys caspica* using MSVAR. *Each line* corresponds to the sum of posterior distributions of the four runs obtained for each cluster; *colour code* as in Fig. 2. *Top* Posterior distribution of the estimated time (in years ago) for each bottleneck. *Bottom* Posterior distribution of ancestral ($N_1$) and present ($N_0$) effective population sizes

highest posterior probabilities may overestimate population structuring (Evanno et al. 2005; Falush et al. 2003). However, when the cluster assignments of the $\Delta K$ method were compared with those based on posterior probabilities, it was obvious that the latter results corresponded largely to a finer substructuring within the clusters of the $\Delta K$ method (Figs. 2, S1).

In the STRUCTURE runs without correction for null alleles, terrapins from Central Anatolia (Fig. 2: cluster A) were clearly distinct from all other terrapins, with very weak indication for any admixture. Terrapins from Eastern Turkey and Syria constituted another cluster (Fig. 2: cluster B). A third distinct cluster corresponded to terrapins from Dagestan (Russia), Azerbaijan and Iran (Fig. 2: cluster C), and the fourth cluster to terrapins from Bahrain and Saudi Arabia (Fig. 2: cluster D). However, four terrapins obtained from the local pet market of Bahrain (Fig. 2: LM) were assigned to cluster B, while terrapins from geographically intermediate sites of clusters B, C and D, and from southernmost Iran, were revealed to have mixed ancestries. In the runs with null alleles coded as recessive, clusters A and B were lumped together and two terrapins with admixed ancestry were referred to this more inclusive cluster A*, but otherwise the results remained unchanged (Table S1).

When demographic history was inferred for the four clusters using MSVAR, an approximately 8,000- to 20,000-year-old bottleneck was suggested for clusters A, B and C. For cluster D a younger bottleneck was inferred that dated only to some 4,000 years ago (Fig. 3; Table 1). Posterior densities of $N_0$, $N_1$ and $T$ were largely consistent over the runs with different starting conditions (Table S3).

*MtDNA*

The 125 cyt *b* sequences corresponded to 14 distinct haplotypes. Ten of these haplotypes had already been identified by Fritz et al. (2008); GenBank accession numbers of the four new haplotypes are HE800433–HE800436. Parsimony network analysis assigned the 14 haplotypes to two distinct clusters. One haplotype cluster is distributed in the west, and the other in the east of the species' range (Fig. 4). The geographical distribution of each haplotype cluster largely matches the composite ranges of two microsatellite clusters for $K$=4 (Fig. 2), and accordingly the western haplotype cluster corresponds to microsatellite cluster A* when null alleles are coded as recessive ($K$=3). Private haplotypes occurred in all microsatellite clusters (Table 2).

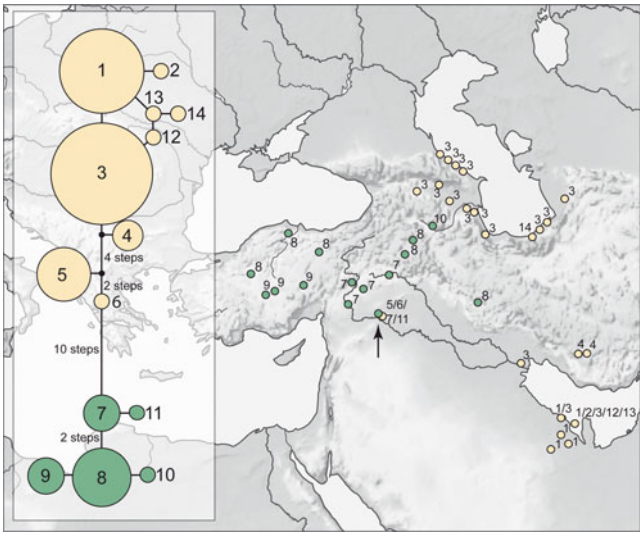


**Fig. 4** Distribution of mitochondrial haplotypes of *Mauremys caspica*. *Green symbols* correspond to the western haplotype cluster; *beige symbols*, to the eastern cluster. The overlapping symbols (*arrow*) represent a syntopic occurrence of haplotypes of the two clusters. *Numbers* indicate individual haplotypes; *numbers separated by slashes*, syntopic occurrences of respective haplotypes. New haplotypes identified in the present study are numbers 11–14. *Inset* Parsimony network showing mutational relationships of haplotypes. *Symbol size* corresponds to haplotype frequency. *Small black circles* are missing node haplotypes. Unless otherwise indicated, *lines connecting haplotypes* represent one mutational step

In the network, the two clusters were separated by a minimum of 10 mutational steps (Fig. 4); within the western cluster, a maximum of 4 steps occurred, and within the eastern cluster, 11 steps. The greatest outgroup probability (0.2796) had haplotype 1, which therefore could be the ancestral haplotype.

At most collection sites not more than one haplotype was found, and only at three localities distinct haplotypes were syntopic (Fig. 4). Six out of the nine haplotypes of the eastern cluster occurred around the Persian Gulf, and at two Gulf localities more than one haplotype was present. At one Syrian site, four haplotypes occurred together, and two of these haplotypes were found only there (haplotypes 5 and 6). This site was also the only one harbouring haplotypes of the western and eastern cluster together.

**Table 1** Demographic inference for *Mauremys caspica* using MSVAR (means for each parameter over the four different runs, cf. Tables S2 and S3)

| Cluster | $N_0$ 50 % (10–90 %) | $N_1$ 50 % (10–90 %) | $T$ (years) 50 % (10–90 %) |
|---|---|---|---|
| A | 398.1 (158.5–1,258.9) | 39,810.7 (12,589.3–100,000.0) | 19,952.6 (6,309.6–79,432.8) |
| B | 1,000.0 (199.5–3,162.3) | 39,810.7 (19,952.6–100,000.0) | 7,943.3 (1,258.9–31,622.8) |
| C | 794.3 (251.2–1,995.3) | 15,848.9 (5,011.9–39,810.7) | 15,848.9 (2,511.9–79,432.8) |
| D | 316.2 (63.1–1,258.9) | 31,622.8 (10,000.0–79,432.8) | 3,981.1 (1,000.0–12,589.3) |

**Table 2** Genetic diversity of STRUCTURE clusters based on 14 microsatellite loci and 1,040 bp of the mitochondrial cyt *b* gene. *n* Number of individuals, $n_A$ number of alleles, $n_{\bar{A}}$ average number of alleles, $n_P$ number of private alleles, *AR* allelic richness, $H_O$ average observed heterozygosity, $H_E$ average expected heterozygosity, $F_{IS}$ average inbreeding coefficient, $n_H$ number of haplotypes (private haplotypes bracketed), *hd* haplotype diversity, *nd* nucleotide diversity. Terrapins with mixed ancestry excluded. All $F_{IS}$ values were statistically significant

| Cluster | *n* | Microsatellites | | | | | | | mtDNA | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $n_A$ | $n_{\bar{A}}$ | $n_P$ | *AR* | $H_O$ | $H_E$ | $F_{IS}$ | $n_H$ | *hd* | *nd* |
| *K*=4 | | | | | | | | | | | |
| A | 16 | 67 | 4.00 | 6 | 4.79 | 0.44 | 0.55 | 0.19 | 2 (1) | 0.592 | 0.0022 |
| B | 22 | 125 | 5.70 | 28 | 8.30 | 0.59 | 0.77 | 0.25 | 6 (5) | 0.724 | 0.0071 |
| C | 41 | 87 | 2.12 | 10 | 4.93 | 0.46 | 0.57 | 0.17 | 3 (1) | 0.096 | 0.0007 |
| D | 38 | 110 | 2.89 | 21 | 6.50 | 0.59 | 0.70 | 0.16 | 4 (3) | 0.324 | 0.0004 |
| *K*=3 | | | | | | | | | | | |
| A* [a] | 40 | 145 | 3.63 | 38 | 10.27 | 0.52 | 0.76 | 0.32 | 8 (7) | 0.817 | 0.0070 |
| C | 41 | 87 | 2.12 | 9 | 7.86 | 0.46 | 0.57 | 0.17 | 3 (1) | 0.096 | 0.0007 |
| D | 38 | 110 | 2.89 | 21 | 6.12 | 0.59 | 0.70 | 0.16 | 4 (3) | 0.324 | 0.0004 |

[a] A*, more inclusive cluster after coding null alleles as recessive. A* embraces clusters A and B plus two terrapins inferred of mixed ancestry under $K = 4$

### Diversity within and divergence among structure clusters

Terrapins of mixed ancestry were excluded from the following analyses. All 14 nuclear microsatellite markers were found to be in linkage equilibrium for each STRUCTURE cluster under *K*=4 and *K*=3. Numbers of alleles per locus ranged from 6 to 19. From a total of 177 alleles, 65 private alleles were found for *K*=4, and 68 private alleles were found for *K*=3 (Table 2).

For *K*=4, clusters B and D had the highest number of private alleles and the highest diversity indices of microsatellites. With respect to mtDNA, cluster B had also the highest haplotype and nucleotide diversities. For *K*=3, the more inclusive cluster A* had most private alleles and the highest diversity values for microsatellites and mitochondrial data.

For *K*=4, average $F_{ST}$ values among the clusters ranged from 0.13 to 0.25 for microsatellites and from 0.53 to 0.97 for cyt *b* sequences (Table 3). For microsatellites, the AMOVA indicated that 22 % of the observed global variation occurred among clusters and 78 % within clusters. For mitochondrial haplotypes, 75 % of the variation occurred among clusters and 25 % within clusters. For *K*=3, fixation indices ranged from 0.12 to 0.25 for microsatellites and from 0.53 to 0.66 for mitochondrial data (Table 4). For microsatellites, the AMOVA found 18 % of the variation to occur among clusters and 82 % within clusters. For mitochondrial data, the respective values were 63 % (among clusters) and 37 % (within clusters).

**Table 3** Pairwise $F_{ST}$ values between STRUCTURE clusters based on microsatellite data (*below diagonal*) and cyt *b* sequences (*above diagonal*) for *K*=4. Individuals with mixed ancestries not considered

| | A | B | C | D |
|---|---|---|---|---|
| A | - | 0.56 | 0.95 | 0.97 |
| B | 0.21 | - | 0.57 | 0.63 |
| C | 0.25 | 0.21 | - | 0.53 |
| D | 0.24 | 0.13 | 0.25 | - |

## Discussion

Our study confirms the earlier finding (Fritz et al. 2008) that *Mauremys caspica* has a weak phylogeographic structure at the mitochondrial level, with one haplotype cluster in the west and another one in the east of its range. By contrast, information from 14 polymorphic microsatellite loci suggests a more complicated pattern with three or four distinct clusters, and we cannot exclude that a denser sampling in central and southern Iran and the inclusion of Iraqi samples will reveal further genetic diversity.

While naïve STRUCTURE analyses resulted in four clusters, only three clusters were suggested when null alleles were coded as recessive. However, the four cluster pattern makes

**Table 4** Pairwise $F_{ST}$ values between STRUCTURE clusters based on microsatellite data (*below diagonal*) and cyt *b* sequences (*above diagonal*) for *K*=3. Individuals with mixed ancestries not considered. Cluster A* includes clusters A and B of Table 3 plus two individuals inferred of mixed ancestry under *K*=4

| | A* | C | D |
|---|---|---|---|
| A* | - | 0.62 | 0.66 |
| C | 0.16 | - | 0.53 |
| D | 0.12 | 0.25 | - |

biological sense. The additional divide between Central Anatolia (cluster A) and the neighbouring populations from Eastern Turkey and Syria (cluster B) coincides with the so-called 'Anatolian diagonal'—a well-known high mountain barrier impeding exchange between western and eastern taxa (Ansell et al. 2011; Davis 1971; Ekim and Güner 1986). Considering the current aggravation of the environmental situation throughout the range of *M. caspica* (Bayazit and Avci 1997; Beaumont 1998; Garstecki and Amr 2011; Harmancioğlu et al. 2001; Hashemi et al. 2012), we suggest therefore that all four microsatellite clusters (Fig. 2) should be treated as distinct management units, pending further sampling in southern and central Iran and in Iraq. The recognition of four management units is also supported by similar $F_{ST}$ values and the demographic scenarios of all four STRUCTURE clusters (Fig. 3; Table 3).

Despite the growing number of studies focusing on the phylogeography of Turkish biota (see the recent review by Ansell et al. 2011), our knowledge of the Near and Middle East is still quite poor. This is especially true for Syria, Saudi Arabia, Iraq and Iran, from where only few studies for freshwater-associated animal species, like fishes (Hrbek and Meyer 2003), frogs (Gvoždík et al. 2010), water snakes (Guicking et al. 2009) or freshwater turtles (Fritz et al. 2008) are available. With respect to Turkey, there is emerging evidence that Central Anatolia harboured during glacial phases open steppe vegetation, providing suitable refuges for temperate animal and plant species (Ansell et al. 2011), as indicated, among others, by the presence of genetically diverse populations of the European pond turtle (*Emys orbicularis*; Fritz et al. 2009) and cluster A of *M. caspica*.

Fritz et al. (2008) hypothesized that outside Central Anatolia two further glacial refugia had existed, namely (1) in the Caucasus and along the southern shore of the Caspian Sea; and (2) all around the Persian Gulf. While our new microsatellite data confirm a distinct genetic entity for the eastern Caucasus region and northern Iran, the pattern around the Gulf is more complicated than previously thought. The terrapins from Saudi Arabia and Bahrain turned out to be clearly different from their conspecifics in southern Iran and the upper courses of Euphrates and Tigris, with evidence for admixture at geographically intermediate collection sites. The distinctiveness of these Arabian terrapins is also supported by the occurrence of private mitochondrial haplotypes (Table 2) and their demographic history (Table 1).

Bottlenecks 4,000–20,000 years ago corroborate that the climatic fluctuations of the Late Pleistocene and Holocene played an important role in shaping the current genetic diversity of *M. caspica* all over the range. The youngest bottleneck, approximately 4,000 years ago, was inferred for Arabian terrapins. This date coincides largely with the onset of the current phase of increasing aridity in the Arabian Peninsula, which commenced about 5,500 years ago after an earlier humid period (Preusser 2009). This event obviously led to a contraction of a formerly larger and more widespread Arabian metapopulation.

All of these lines of evidence clearly contradict previous speculations that the Arabian populations, confined to oasis complexes, are introduced (Gasperetti et al. 1993). Four pet trade terrapins obtained in Bahrain, however, were clearly distinct from the native Arabian populations, and were assigned to the Euphrates-Tigris cluster B by microsatellite data (Fig. 2). This situation highlights the danger of genetic pollution, which is particularly imminent in the endangered Arabian populations. Red-eared sliders (*Trachemys scripta elegans*) and cichlids have been already introduced into their habitats (Anon. 2003), suggesting that the release of non-native *M. caspica* is only a question of time.

Moreover, our results are in accord with the conclusion of Fritz et al. (2008) that genetic data do not support the validity of any of the three morphologically defined subspecies of *M. caspica*. Fritz and Wischuf (1997) distinguished a northern and two southern subspecies but their distribution ranges conflict both with the geographical distribution of mitochondrial haplotype clusters and of the microsatellite clusters (compare the map of Fig. 1 in Fritz et al. 2008 with our Figs. 2 and 4), which is why we propose to abandon their usage.

**Acknowledgments** Thanks for support go to Matjaž Gregorič, Michael Mende, Anke Müller, Ole Reuter, and Mario Vargas-Ramírez. Markus Auer, Daniel Frynta, Mario Herz and Pavel Široký donated samples. We are indebted to His Highness Sheikh Dr. Sultan bin Mohammed Al Qasimi, Supreme Council Member and Ruler of Sharjah, and Hana Al Suwaidi, Chief Executive Officer of EPAA, Sharjah, for the opportunity to study Arabian terrapins. M.V. was funded by a PhD fellowship of the German Academic Exchange Service (DAAD; A/09/91179). Laboratory work was funded by the German Research Foundation (DFG; FR 1435/7-1) and field work in Turkey by the Scientific and Technological Research Council of Turkey (TÜBITAK; TBAG-2402 103 T189).

## References

Anderson, S. C. (1979). Synopsis of the turtles, crocodiles, and amphisbaenians of Iran. *Proceedings of the California Academy of Sciences, 4th Series, 41*, 501–528.

Anon. (2003). *Conservation Assessment and Management Plan (CAMP) for the Fauna of Arabia, 23-26 February 2003. Fourth International Conservation Workshop for the Threatened Fauna of Arabia. Final Report*. Sharjah: Breeding Center for Endangered Arabian Wildlife.

Ansell, S. W., Stenøien, H. K., Grundmann, M., Russell, S. J., Koch, M. A., Schneider, H., et al. (2011). The importance of Anatolian mountains as the cradle of global diversity in *Arabis alpina*, a key arctic-alpine species. *Annals of Botany, 108*, 241–252.

Arctander, P. (1988). Comparative studies of avian DNA by restriction fragment polymorphism analysis. *Journal of Ornithology, 129*, 205–216.

Bannikov, A. G. (1951). Materialy k poznaniyu biologii kavkazskikh cherepakh. *Uchenye zapiski Moskovskogo gorodskogo pedagogicheskogo instituta W. P. Potemkina, 18*, 129–167.

Bayazit, M., & Avci, I. (1997). Water resources of Turkey: potential, planning, development and management. *International Journal of Water Resources Development, 13*, 443–452.

Beaumont, P. (1998). Restructuring of water usage in the Tigris-Euphrates basin: The impact of modern water management policies. In J. Albert, M. Bernhardsson, & R. Kenna (Eds.), *Transformations of Middle Eastern Natural Environments: Legacies and Lessons* (pp. 168–186). New Haven: Yale School of Forestry and Environmental Studies.

Beaumont, M. A. (1999). Detecting population expansion and decline using microsatellites. *Genetics, 153*, 2013–2029.

Biricik, M., & Turğa, Ş. (2011). Description of an Euphrates softshell turtle (*Rafetus euphraticus*) nest from the Tigris River (SE Turkey). *Salamandra, 47*, 99–102.

Chikhi, L., Sousa, V., Luisi, P., Goossens, B., & Beaumont, M. A. (2010). The confounding effects of population structure, genetic diversity and the sampling scheme on the detection and quantification of population size changes. *Genetics, 186*, 983–995.

Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology, 9*, 1657–1660.

Davis, P. H. (1971). Distribution patterns in Anatolia with particular reference to endemism. In P. H. Davis, P. C. Harper, & I. C. Hedge (Eds.), *Plant Life of South-West Asia* (pp. 15–27). Edinburgh: Royal Botanic Garden.

Ekim, T., & Güner, A. (1986). The Anatolian diagonal: fact or fiction? *Proceedings of the Royal Society of Edinburgh, 89B*, 69–77.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology, 14*, 2611–2620.

Excoffier, L., Laval, G., & Schneider, S. (2005). ARLEQUIN (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online, 1*, 47–50.

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics, 164*, 1567–1587.

Falush, D., Stephens, M., & Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes, 7*, 574–578.

Fritz, U., & Havaš, P. (2007). Checklist of chelonians of the world. *Vertebrate Zoology, 57*, 149–368.

Fritz, U., & Wischuf, T. (1997). Zur Systematik westasiatisch-südosteuropäischer Bachschildkröten (Gattung *Mauremys*). *Zoologische Abhandlungen, Staatliches Museum für Tierkunde Dresden, 49*, 223–260.

Fritz, U., Auer, M., Bertolero, A., Cheylan, M., Fattizzo, T., Hundsdörfer, A. K., et al. (2006). A rangewide phylogeography of Hermann's tortoise, *Testudo hermanni* (Reptilia: Testudines: Testudinidae): implications for taxonomy. *Zoologica Scripta, 35*, 531–543.

Fritz, U., Ayaz, D., Buschbom, J., Kami, H. G., Mazanaeva, L. F., Aloufi, A. A., et al. (2008). Go east: phylogeographies of *Mauremys caspica* and *M. rivulata* – discordance of morphology, mitochondrial and nuclear genomic markers and rare hybridization. *Journal of Evolutionary Biology, 21*, 527–540.

Fritz, U., Ayaz, D., Hundsdörfer, A. K., Kotenko, T., Guicking, D., Wink, M., et al. (2009). Mitochondrial diversity of European pond turtles (*Emys orbicularis*) in Anatolia and the Ponto-Caspian Region: multiple old refuges, hotspot of extant diversification and critically endangered endemics. *Organisms, Diversity and Evolution, 9*, 100–114.

Garstecki, T., & Amr, Z. (2011). *Biodiversity and Ecosystem Management in the Iraqi Marshlands—Screening Study on Potential World Heritage Nomination*. Amman: IUCN.

Gasperetti, J., Stimson, A. F., Miller, J. D., Ross, J. P., & Gasperetti, P. R. (1993). Turtles of Arabia. *Fauna of Saudi Arabia, 13*, 170–367.

Glaubitz, J. C. (2004). CONVERT: a user friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Molecular Ecology Notes, 4*, 309–310.

Gmelin, S. G. (1774). *Reise durch Rußland zur Untersuchung der drey Natur-Reiche. Dritter Theil. Reise durch das nordliche Persien, in den Jahren 1770, 1771, bis im April 1772*. St. Petersburg: Kayserliche Academie der Wissenschaften.

Goudet, J. (1995). FSTAT (version 1.2): a computer program to calculate *F*-statistics. *Journal of Heredity, 86*, 485–486.

Guicking, D., Joger, U., & Wink, M. (2009). Cryptic diversity in a Eurasian watersnake (*Natrix tessellata*, Serpentes: Colubridae): evidence from mitochondrial sequence data and nuclear ISSR-PCR fingerprinting. *Organisms, Diversity and Evolution, 9*, 201–214.

Gvoždík, V., Moravec, J., Klütsch, C., & Kotlík, P. (2010). Phylogeography of the Middle Eastern tree frogs (*Hyla*, Hylidae, Amphibia) as inferred from nuclear and mitochondrial DNA variation, with a description of a new species. *Molecular Phylogenetics and Evolution, 55*, 1146–1166.

Hall, T. A. (1999). BIOEDIT: a user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series, 41*, 95–98.

Harmancioğlu, N., Alpaslan, N., & Boelee, E. (2001). *Irrigation, health and environment: a review of literature from Turkey*. Colombo: International Water Management Institute (IWMI).

Hashemi, S. A. R., Eskandary, G. H., & Ansary, H. (2012). Biomass of fish species in the Shadegan wetland, Iran. *Research Journal of Recent Sciences, 1*, 66–68.

Hrbek, T., & Meyer, A. (2003). Closing of the Tethys Sea and the phylogeny of Eurasian killifishes (Cyprinodontiformes: Cyprinodontidae). *Journal of Evolutionary Biology, 16*, 17–36.

Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources, 9*, 1322–1332.

Kinzelbach, R. (1986). Recent records of the Nile softshell turtle, *Trionyx triunguis*, and of the Euphrates softshell turtle, *Trionyx euphraticus*, in the Middle East. *Zoology in the Middle East, 1*, 83–87.

Librado, P., & Rozas, J. (2009). DNASP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics, 25*, 1451–1452.

Palsbøll, P. J., Bérubé, M., & Allendorf, F. W. (2011). Identification of management units using population genetic data. *Trends in Ecology & Evolution, 22*, 11–16.

Preusser, F. (2009). Chronology of the impact of Quaternary climate change on continental environments in the Arabian Peninsula. *Comptes Rendus Geoscience, 341*, 621–632.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*, 945–959.

Randi, E. (2008). Detecting hybridization between wild species and their domesticated relatives. *Molecular Ecology, 17*, 285–293.

Rousset, F. (2008). GENEPOP'007: a complete reimplementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources, 8*, 103–106.

Spinks, P. Q., Shaffer, H. B., Iverson, J. B., & McCord, W. P. (2004). Phylogenetic hypotheses for the turtle family Geoemydidae. *Molecular Phylogenetics and Evolution, 32*, 164–182.

Storz, J. F., & Beaumont, M. A. (2002). Testing for genetic evidence of population expansion and contraction: an empirical analysis of

microsatellite DNA variation using a hierarchical Bayesian model. *Evolution, 56*, 156–166.

Vamberger, M., Stuckas, H., & Fritz, U. (2011). Fifteen microsatellite markers for the stripe-necked terrapin *Mauremys caspica* (Testudines: Geoemydidae) and cross-amplification tests in *M. rivulata*. *Conservation Genetics Resources, 3*, 87–89.

van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes, 4*, 535–538.

Weir, B. S., & Cockerham, C. C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution, 38*, 1358–1370.